Friendship-Association Based Anonymous Identical Users Detection System in Multiple SMNs

[1]Maske Shreyas Tukaram, [2]B.R.Solunke

[1]ME Student, [2]Professor
[1,2] Computer Science and Engineering,
[1,2] N.B. Nawale Sinhgad College of Engineering, Solapur

***Abstract*: Somewhat recently, many sorts of interpersonal interaction locales have arisen and contributed massively to huge volumes of certifiable information on friendly practices. Perceiving unknown, yet indistinguishable clients among numerous SMNs is as yet an obstinate issue. Obviously, cross-stage investigation might assist with tackling numerous issues in friendly processing in both hypothesis and applications. Since the public profiles can be effortlessly copied and handily imitated by clients with various purposes, most current client recognizable proof goals, which basically center around text mining of clients' public profiles, are delicate. A few investigations have endeavored to coordinate with clients dependent on the area and timing of client content. In addition, since online SMNs are very symmetric, existing client ID plans dependent on network structure are not compelling. This present reality companion cycle is profoundly individual and essentially no two clients share a compatible companion cycle. Thusly, it is more exact to utilize a fellowship construction to dissect cross-stage SMNs. Since indistinguishable clients will in general set up halfway comparative fellowship structures in various SMNs. we proposed the Friend Association-Based User Identification (FAUI) calculation. FAUI works out a coordinating with degree for all up-and-comer User Matched Pairs (UMPs), and just UMPs with highest levels are viewed as indistinguishable clients. We likewise foster two suggestions to work on the proficiency of the calculation. Consequences of broad examinations show that FAUI performs far superior to current organization structure-based calculations. We additionally proposed the composing style measurements. Which utilizes Capitalization, normal word length, accentuation, utilization of images, utilization of numerals, dividing, new lines/sections, short structures, utilization of shoptalk, utilization of smileys, common words.**

***Index Terms*: Cross-Platform, Social Media Network, Anonymous Identical Users, Friend Relationship, User Identification**

I. INTRODUCTION

The most recent couple of years have seen the development and advancement of an energetic examination stream on an enormous assortment of online Social Media Network (SMN) stages. Twitter is the biggest miniature blog administration, has in excess of 600 million clients and creates as much as 340 million tweets each day. Because of this variety of online web-based media organizations, individuals will in general utilize diverse Social Media Networks for various purposes. As such, every existent Social Media Network fulfills some client needs. As far as Social Media Network the executives, coordinating with unknown clients across various Social Media Network stages can give incorporated subtleties on every client and educate relating guidelines, for example, focusing on administrations arrangements.

Notwithstanding, practically all new Social Media Network - put together examinations center with respect to a solitary Social Media Network stage, yielding inadequate information. Consequently, this review researches the methodology of intersection various Social Media Network stages to illustrate these practices. All in all, every existent Social Media Network fulfills some client needs. Cross-stage research faces various difficulties. Cross-stage approach has blended different Social Media Network stages to make more extravagant crude information and more complete Social Media Networks for social processing undertakings.

Client recognizable proof means client acknowledgment, client personality goal, client coordinating, and anchor connecting. Albeit no arrangement can distinguish all indistinguishable unknown Social Media Network clients, some Social Media Network components might be utilized to recognize a part of clients across numerous Social Media Networks. Many examinations have tended to the client distinguishing proof issue by looking at public client profile ascribes, including screen name, birth-day, area, sexual orientation, profile photograph, and so forth Since these characteristics don't need eliteness and are effortlessly faked by clients for various purposes. A few analysts have utilized public client exercises to perceive clients utilizing post time, area. Since area information is hard to acquire.

Online Media Network associations fall into two classifications: single-following associations and shared after associations. Single-following associations are additionally called following connection ships or following connections. On the off chance that client X follows client Y, client X and client Y have an after relationship(single-way fans in which one knows the other, yet not the other way around). Following connections are normal in miniature contributing to a blog Social Media Networks, such Twitter Micro blog. In like manner, shared after associations are called companion connections. In miniature publishing content to a blog SMNs, a companion relationship alludes to the common after connections between two clients. In most other SMNs, for example, Facebook and We-visit, a companion relationship shapes just if a companion demand is sent by one client and affirmed by the other client. Companion connection ships are hard to counterfeit by malignant clients, and along these lines reflect certifiable connections much better.

Regardless, cross-stage research faces various difficulties. As displayed in Fig. 1, with the development of SMN stages on the Internet, the cross-stage approach has blended different SMN stages to make more extravagant crude information and more complete SMNs for social processing undertakings. SMN clients structure the regular scaffolds for these SMN stages. The essential theme for cross-stage SMN research is client recognizable proof for various SMNs. Investigation of this theme establishes a framework for additional cross-stage SMN research.

II. Related Work

Perito et al. [3] determined the likeness of screen names and recognized clients utilizing paired classifiers. A few investigations tending to mysterious client recognizable proof have zeroed in on open profile ascribes, including screen name, sex, birthday, city, and profile picture. A screen name is the publically required profile include in practically all SMNs. It has been generally investigated as an approach to perceive clients across various SMNs.

Additionally, Liu et al. [4] coordinated with clients in a solo methodology utilizing screen names. Zafarani and Liu [5] proposed a technique to plan personalities across various SMN stages, exactly approving a few theories. On top of this work, they [6] further fostered a client planning strategy by demonstrating client conduct on-screen names. Among public profile ascribes, the profile picture is another component that has gotten impressive review. Acquisti et al. [7] tended to the client distinguishing proof assignment with a face acknowledgment calculation. Albeit both screen names and profile pictures can recognize clients, they can't be applied to huge SMEs. This is on the grounds that a few clients might have a similar screen name and profile pictures.

For instance, numerous clients have the screen name "John Smith" on Facebook. Clearly, utilizing a blend of profile provisions can bring about better client ID. Tofu et al. [8] proposed a methodology by estimating the distance between client profiles. Motoyama and Varghese [9] accumulated characteristics (training, occupation, and so on) resources of words and coordinated with clients by computing the comparability of clients. Goga [10] connected records having a place with a similar individual's personality, in light of on the profile data. Cortis [11] proposed a weighted philosophy based client profile goal strategy. Abel et al. [12] collected client profiles and coordinated with clients across frameworks. Comparative investigations across numerous stages are additionally found in [13], [14], [15]. Without a doubt, public profile ascribes give amazing data to client ID. Be that as it may, a few credits are copied in enormous scope SMNs and are effectively mimicked. In this manner, simply profile-based plans have limits when they are applied to enormous scope SMNs.

Content-Based User Identification arrangements endeavor to perceive clients dependent on the occasions and areas that clients post substance, just as the composing style of the substance. Zheng et al. [18] proposed a structure for creation distinguishing proof utilizing the composing style of online messages and characterization methods. Almishari and Tsudik [19] proposed connecting clients across various SMNs by taking advantage of the composing style of creators. Kong and Zhang [20] proposed Multi-Network Anchoring (MNA) to plan clients. They determined the joined similitudes of client's social, spatial, transient, and text data in various SMNs, and analyzed a stable coordinating with issue between two arrangements of client accounts.

Goga et al. [21] took advantage of the geo-area connected to clients' posts, the timestamp of posts, and clients' composing style to address client ID errands. Geo-area seems to have strong elements for client acknowledgment. Notwithstanding, this data is frequently inadequate in SMNs, since just a little part of clients will post their areas. In spite of the fact that composing style arrangements perform well in situations including long substance, these procedures are not pertinent to SMNs like Twitter and Sina Microblog, in which short sentences are in all probability posted. 2.3 Network Structure-Based User Identification Network structure-based examinations on client ID across various SMNs are utilized to perceive indistinguishable clients exclusively by client network designs and seed, or priori, distinguished clients. As displayed above, network-based client distinguishing proof represents a few significant difficulties, with few examinations to expand . To resolve this issue,

Bartunov et al. [23] coordinated profiles with a net work structure utilizing a Conditional Random Fields model and acquired better client ID results. Organization structure-based client distinguishing proof is a quite a problem and can be utilized to recognize just a piece of indistinguishable clients. NS, the principal network structure-based client acknowledgment calculation across SMNs, can complete client acknowledgment undertakings by utilizing just the organization structure and recognized 30.8% indistinguishable clients in a ground-truth dataset [22]. Assume that there are two SMNs: SMNA and NB. NS initially ascertains a bunch of planning scores for each single, unmapped client substance in SMNA to each unmapped client element in NB. Then, at that point a flightiness is applied to decide if a client in SMNB can be coordinated. Just if the flightiness is bigger than a limit would a client match be acknowledged. Furthermore, NS requires an opposite match to affirm the client match, which is exorbitant in tests. JLA endeavors to coordinate with unmapped hubs from various charts by contrasting the planned neighbors of every hub. It ascertains an organization distance between any two unmapped hubs in two diverse undirected organizations.

To break down security and namelessness, Narayanan and Shmatikov [23] created NS, in light of on network geography. Like FRUI, NS and JLA are coordinated guides. To accommodate the SMNs, Korula, and Lattanzi [24] introduced a many-to-many planning calculation dependent on the levels of unmapped clients and the quantity of normal neighbors, utilizing two control boundaries to adjust execution. These works had a comparative work process, discovering seed clients first, then, at that point utilizing these seed clients to recursively spread data through networks and broaden sets of planned hubs. The undertaking on client distinguishing proof is firmly identified with the de-anonymization issue for protection safeguarding informal community

examination, which re-recognizes people in online distributed SMN datasets. In this specific situation, SMN information are anonymized before discharge. Zhou and Pei broke down the local assaults of de-anonymization and proposed security safeguarding approaches utilizing k-namelessness and l variety. Other de-anonymization assaults have additionally been dissected. Since a cross-stage client distinguishing proof is like the de-anonymization task, it tends to be applied to address the de-anonymization issue.

Organic product performs obviously superior to NS, the de-anonymization calculation. The joint utilization of profile data, client practices concealed cave content and organization constructions might prompt better outcomes. Jain and Kumaraguru [16], [17] created Finding Nemo, a strategy that matches Facebook and Twitter accounts. Be that as it may, this text-based organization search technique has low precision and high intricacy as far as client ID, since just the texts of similar epithets are perceived while looking through the companion sets of companions [12, 13]. Additionally, exact examinations show that specific profiles can be coordinated dependent on the organization structure utilizing JLA. In this review, we propose an imaginative way to deal with address the difficulties looked by past investigations. This new methodology centers around the fellowship structure and fosters the Friend Relationship-based User Identification (FRUI) calculation. Natural product varies from the two existing calculations, JLA and NS, in the accompanying angles (see Table 1): (1) NS is appropriate for coordinated organizations, while JLA and FRUI center around undirected organizations. JLA is limited in undirected organizations by Conditional Random Fields, while FRUI depends on companion connections, as this is more solid and predictable with genuine fellowships. (2) NS requires an extra control boundary (flightiness edge) to distinguish client matches. In the event that the unpredictability is over still up in the air edge, NS acknowledges an applicant User Matched Pair (UMP). Unmistakably, the limit is a free boundary and ought to be given ahead of time. Conversely, no additional free boundaries are needed by JLA and FRUIT. (3) JLA analyzes unmapped neighbors of hubs from one of the two SMNs, while NS matches unidentified clients from various organizations by looking at the planned neighbors of every hub. Organic product plans to distinguish the most coordinated with sets among planned clients yet doesn't emphasize unmapped clients. Accordingly, it extraordinarily diminishes computational intricacy. (4) NS utilizes unmapped clients' in-and out-degrees, just as the recognized clients, to work out scores in coordinated organizations. JLA, conversely, utilizes recognized clients and their certifications. Any planned client has various degrees in various SMNs.

Hence, subtleties on how these degrees are gotten ought to be talked about ahead of time. Similarly, just distinguished clients are needed in our FRUIT. (5) NS processes the coordinating with degree by shared known active/approaching neighbors and out-/in-degrees, with the suspicion that clients in various SMNs have comparative active/approaching neighbors. JLA utilizes a dice coefficient to work out the coordinating with degree, which might confound clients with a high likelihood when two clients share a couple of known companions. Interestingly, FRUI considers the quantity of shared known companions in match degree computation.

We propose this method because: (1) Important information is usually organized in list formats by websites. They may repeatedly occur in a sentence that is separated by commas, or be placed side by side in a well-formatted structure (e.g., a table). This is caused by the conventions of webpage design. Listing is a graceful way to show parallel knowledge or items and is thus frequently used by webmasters. (2) Important lists are commonly supported by relevant websites and they repeat in the top search results, whereas unimportant lists just infrequently appear in results. This makes it possible to distinguish good lists from bad ones, and to further rank facets in terms of importance. Experimental results confirm the above observations and demonstrate that the query facets mined by aggregating them are meaningful.

III. PROPOSE SYSTEM

To introduce a uniform arrangement system for unknown indistinguishable client discovery by utilizing network construction and composing style measurements and characterizes the issue of companion affiliation based client planning. The major objective is finding query facets which are various social occasions of words or articulations that explain and summarize the substance covered by query. To give diverse social event of query facets are explicitly useful for unclear requests. To offer direct information or second responses that customers are searching for.

To deal with the assortment of the best ten results. We can re-rank yield to do whatever it takes not to show the pages that are near replicated in query facets at the top. To distinguish comparative clients in distinction social stages utilizing network structure. To recognize comparative clients in contrast social stages utilizing composing style measurements. To execute the calculation to work on the productivity of the proposed work. To show that it gives preferable execution over current organization structure-based calculation. Proposed approach is utilized recognize mysterious indistinguishable clients in different Social Media Networks. The recognizable proof done in different Social Media Network utilizing network design and utilizing composing style measurements. This will assist specialists with gathering client profile information from different SMN's and relate it for better client conduct investigation. The proposed approach for ID of indistinguishable clients in numerous Social-Media Network utilizing composing style measurements dependent on after focuses: Use of Capitalization, normal word length, accentuation, images, numerals, dispersing, new lines or passages, short structures, smileys, run of the mill words. The Friend Association-Based User Identification (FAUI) calculation. FAUI computes a match degree for all applicant User Matched Pairs (UMPs), and just UMPs with highest levels are considered as indistinguishable clients. We likewise created two recommendations to work on the productivity of the calculation. Aftereffects of broad investigations exhibit that FAUI performs far superior to current organization structure-based calculations.

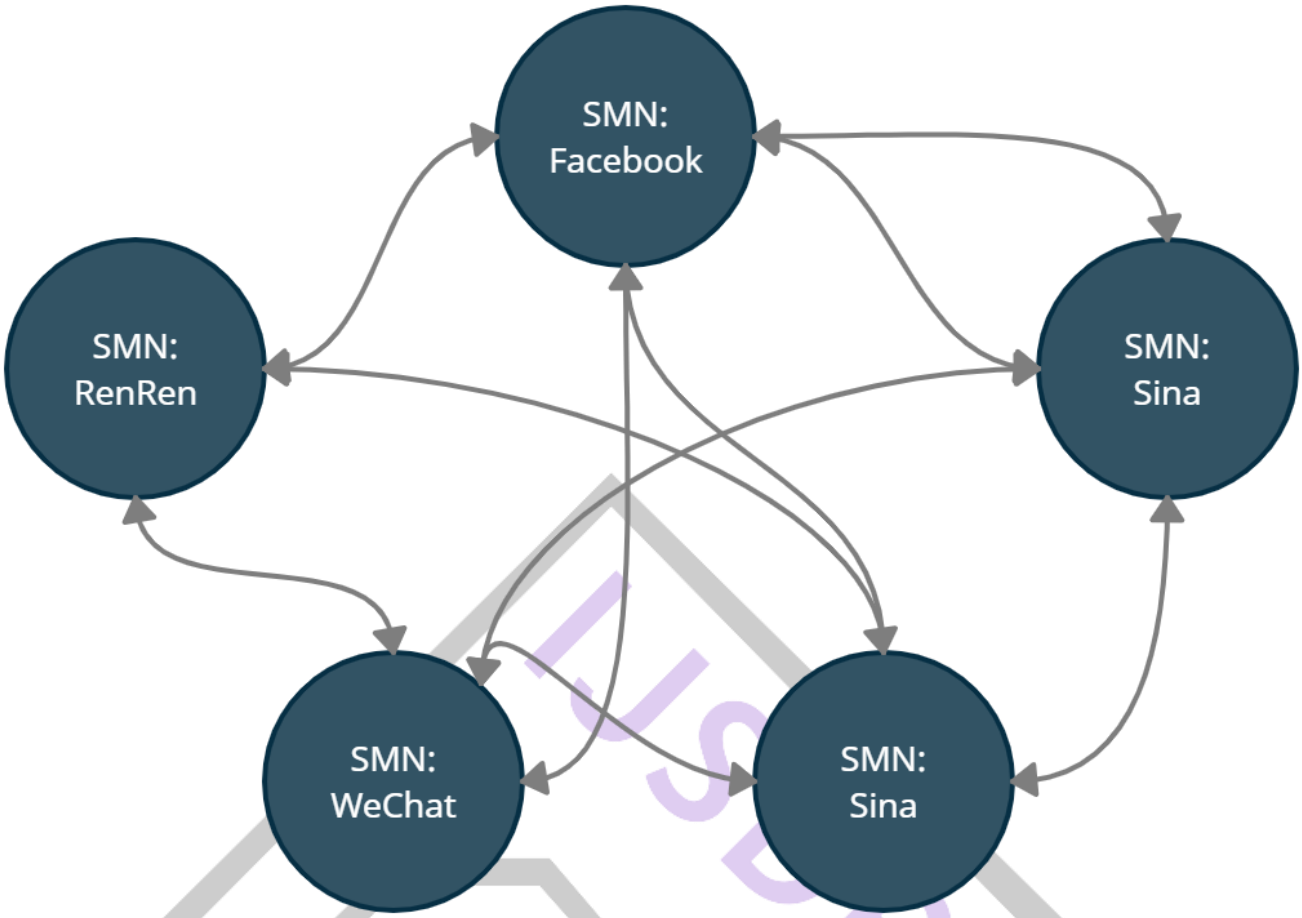

Figure 1 Block Diagram of Proposed System

In reality, we can gather that every individual has his own companion cycle, which is exceptionally person. Accordingly, in the event that we know the entirety of an individual's companions, we presumably know what his identity is. Utilizing SMNA in Fig. 3(a) for instance, on the off chance that one has just client 1 as a companion, clearly he should be client 3. In the event that somebody asks who has the companion set of clients 1 and 2, it is clearly client 4. Clients will in general have comparable companions across various SMNs

Thusly, we can estimate that: (1) If some Valid Priori UMPs are given, then, at that point a bunch of applicant UMPs can be derived, and (2) the more realized companions are partaken in a competitor UMP, the higher the likelihood that they have a place with a similar person. Utilizing the completely covered SMNA and SMNB, as a delineation, Fig. 3(c) shows SMNA and SMNB with the Priori UMPs UMPA~B(1, 1) and UMPA~B(2, 2). Naturally, (UEA1, UEB1) and (UEA2, UEB2) are put together. Then, at that point we find that UEA4 and UEB4 share a similar companion set, which is the larg est set dependent on the current UMP set. We would then be able to presume that UEA4 and UEB4 have a decent potential for success of framing another UMP. After UMPA~B(4, 4) is recognized, it is added to Priori UMPs. By more than once utilizing the above technique, UMPA~B(5, 5), UMPA~B(6, 6), UMPA~B(7, 7), and UMPA~B(1, 1) are iden tified sequentially.

A preprocessor is intended to procure whatever number Priori UMPs as would be prudent. At present, there is no normal ap proach accessible to acquire UMPs between two SMNs. Indicated techniques should be planned by given SMNs. Albeit no bound together interaction is appropriate for the Pre processor, a few calculations can be embraced by the application, e.g., email address, screen name, URL, and so forth An email address seems, by all accounts, to be an interesting element for each record, and can be utilized to gather Priori UMPs. Bal duzzi et al. [38] investigated email locations to discover indistinguishable clients among various SMNs with the "Companion Finder" system. Notwithstanding, since email addresses are private, virtually all SMNs have incapacitated the "Companion Finder".

Algorithm 1: Friendship-Association Based Anonymous Identical Users Identification.

Input: SMNA, SMNB, Priori UMPs: PUMPs
Output: Identified UMPs: UMPs
1:function FAUI(SMNA, SMNB, PUMPs)
2:T = {}, R = dict(), S = PUMPs, L = [], max = 0, FA = [], FB = []

3:while S is not empty do
4: Add S to T

5: if max > 0 do
6: Remove S from L[max]
7: while L[max] is empty
8: max = max – 1
9: if max == 0 do
10: return UMPs
11: Remove UMPs with mapped UE from L[max]
12: foreach UMPA-B(i, j) in S do
13: foreach UEAain the unmapped neighbors of UEAido
14: FA[i] = FA[i] + 1
15:foreachUEAbin the unmapped neighbors of UEAjdo
16: R[UMPA~B(a, b)] += 1, FB[j] = FB[j] + 1
17: Add UMPA~B(a, b) to L[R[UMPA~B(a, b)]]
18: if R[UMPA~B(a, b)] > max do
19: max = R[UMPA~B(a, b)]
20: m = max, S = {}
21: while S is empty do
22: Remove UMPs with mapped UE from L[max]
23: C = L[m], m = m - 1, n = 0
24: S = {un-Controversial UMPs in C }
25:while S is empty do
26: n = n + 1, I = {UMPs with top n Mijin C using (5)}
27:S = {un-Controversial UMPs in I }
28: if I=c do
29: break
30: return T

In the implementation, the Identifier first calculates matrix R using Proposition 1 and initializes the match degree. Then it iterates and identifies UMPs using function g until no UMP can be identified. In each iteration, once the UMPs are identified, the items are removed from the Candidate UMP list, and R is recalculated based on Proposition2 . The process is summarized in Algorithm .

Proposition 1:Given two SMNs, SMNA and SMNB, with s pairs of Priori UMPs, them× n matrix R= QAPB contains the numbers of shared known friends, where m and n are the numbers of Adjacent Users in SMNA and SMNB, rij stands for the number of shared friends of UIAI and UIBj in the s pairs of UMP, QA and PB denote the connections between Adjacent Users and identified users in SMNA and the connections between identified users and Adjacent Users in SMNB, respectively.

Proposition 2:In the t-iteration process, if k UMPs are generated then, R(t+1)=Combine(R(t), Delta R) Where the function combine removes the items with any UE included in the k UMPs, and returns the union of the remaining ones in R(t) and Delta R. The union operation adds the value of the items in both R(t) an Delta R, and joins the left items.

Priori User Match Pairs: Priori UMPs are UMPs given in advance, before user identification resolution work is executed. Priori UMPs are often used as the condition to identify more User Match Pairs.

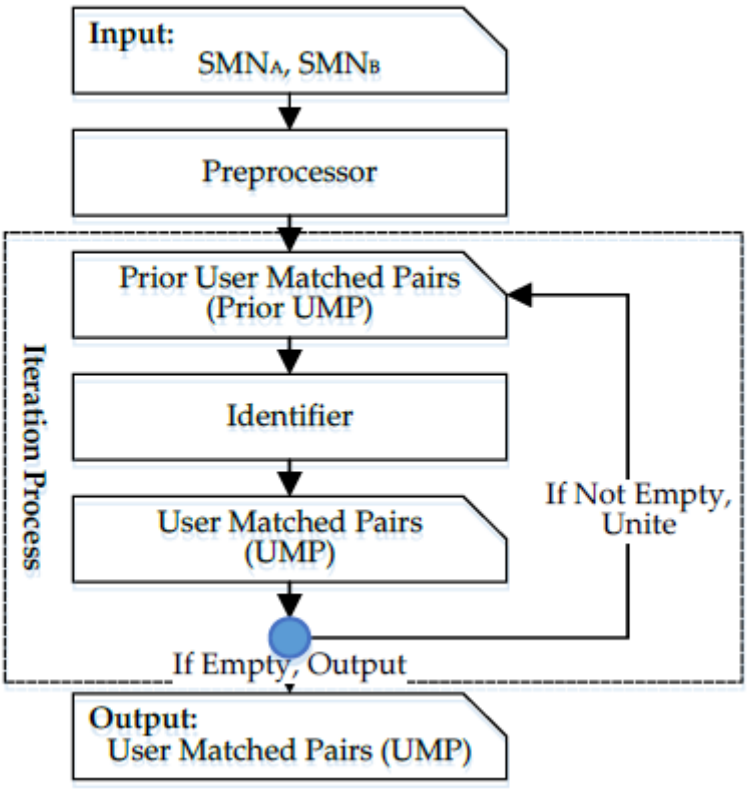

Figure 2: Uniform solution framework.

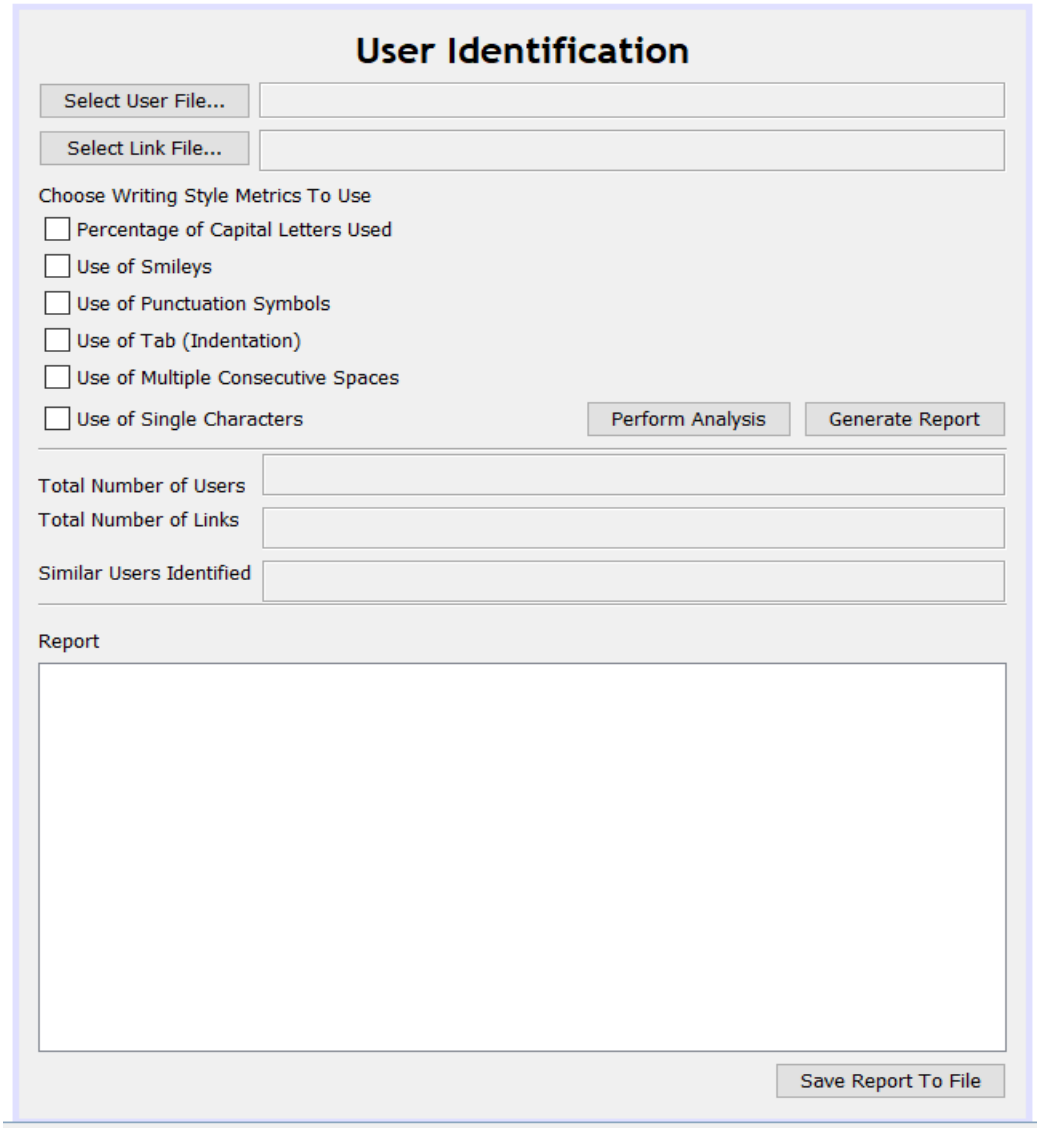

Figure 2 User Identification

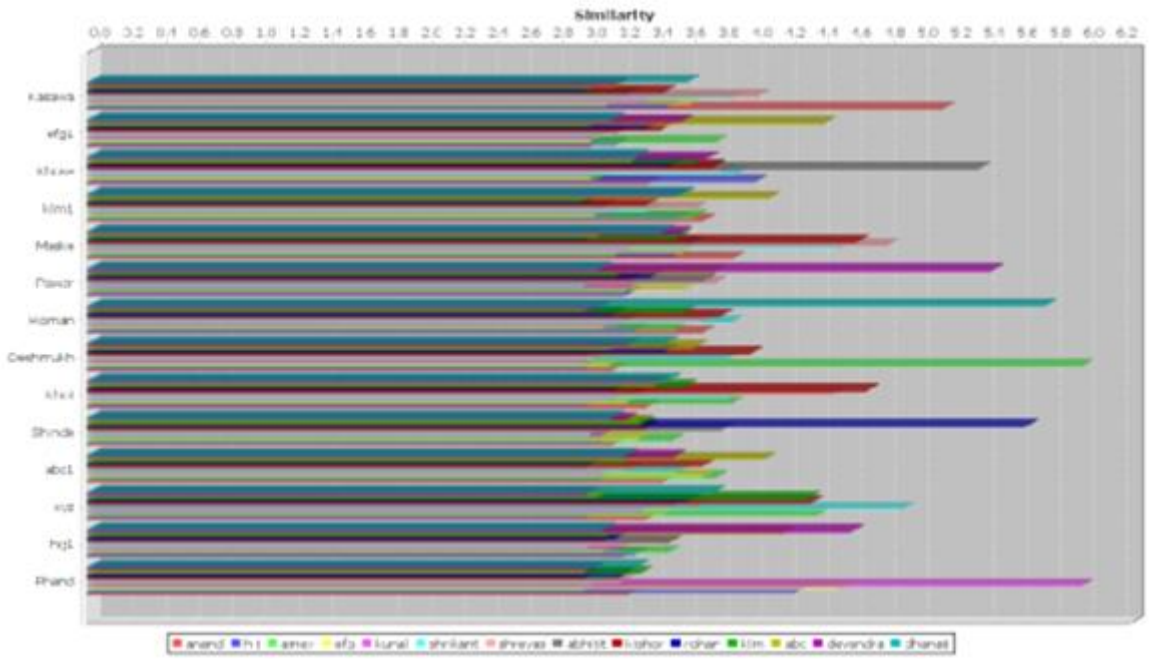

Figure 3 Similarity Graph

This review resolved the issue of client ID across SMN stages and offered an inventive arrangement. As a vital part of SMN, network structure is of foremost significance and helps settle deanonymization client ID assignments. Thusly, we proposed a uniform organization structure-based client recognizable proof arrangement. We likewise fostered a clever companion relationship-based calculation called FRUI. To work on the productivity of FRUI, we de-scribed two suggestions and tended to the intricacy. At last, we confirmed our calculation in both manufactured net-works and ground-truth organizations.

Aftereffects of our exact investigations uncover that organization construction can achieve significant client ID work. Our FRUI calculation is straightforward, yet effective, and performed far superior to NS, the current condition of-craftsmanship network structure-based client ID arrangement. In situations when crude text information is inadequate, fragmented, or difficult to acquire because of protection settings, FRUI is incredibly reasonable for cross-stage assignments.

IV. Conclusion

In this paper, In various web-based media organizations (SMNs), this present reality companion cycle is exceptionally individual and basically no two clients share a harmonious companion cycle. In this manner, it is more exact to utilize a kinship construction to dissect cross-stage SMNs. Since indistinguishable clients will in general set up fractional comparative kinship structures in various SMNs. In this thesis, we proposed the Friend Association-Based User Identification (FAUI) calculation. FAUI computes a match degree for all applicant User Matched Pairs (UMPs), and just UMPs with highest levels are considered as indistinguishable clients. We additionally created two recommendations to work on the productivity of the calculation. Consequences of broad investigations show that FAUI performs obviously superior to current organization structure-based calculations. In this thesis, we resolved the issue of client recognizable proof across Social Media Network stages and offered a creative arrangement. As a critical part of Social Media Network, network structure is of fundamental significance and helps settle deanonymization client recognizable proof assignments. Recognizing unknown clients across different Social Media Networks is testing work. Consequently, just a piece of indistinguishable clients with various monikers can be perceived with this technique.

References

[1] Wikipedia, "Twitter, " http://en.wikipedia.org/wiki/Twitter. 2014.

[2] Xinhuanet, "Sina Microblog Achieves over 500 Million Users," http://news.xinhuanet.com/tech/2012-02/29/c_122769084.htm. 2014.

[3] D. Perito, C. Castelluccia, M.A. Kaafar, and P. Manils, "How unique and traceable are usernames?," Privacy Enhancing Technologies (PETS'11), pp. 1-17, 2011.

[4] J. Liu, F. Zhang, X. Song, Y.I. Song, C.Y. Lin, and H.W. Hon, "What's in a name?: an unsupervised approach to link users across communities," Proc. of the 6th ACM international conference on Web search and data mining(WDM'13), pp. 495-504, 2013.

[5] R. Zafarani and H. Liu, "Connecting corresponding identities across communities," Proc. of the 3rd International ICWSM Conference, pp. 354-357, 2009.

[6] R. Zafarani and H. Liu, "Connecting users across social media sites: a behavioral-modeling approach, " Proc. of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'13), pp.41-49, 2013.

[7] A. Acquisti, R. Gross and F. Stutzman, "Privacy in the age of augmented reality," Proc. National Academy of Sciences, 2011.

[8] T. Iofciu, P. Fankhauser, F. Abel, and K. Bischoff, "Identifying users across social tagging systems," Proc. of the 5th International AAAI Conference on Weblogs and Social Media, pp. 522-525, 2011.

[9] M. Motoyama and G. Varghese, "I seek you: searching and matching individuals in social networks," Proc. of the 11th international workshop on Web Information and Data Management (WIDM'09), pp. 67-75, 2009.

[10] O. Goga, D. Perito, H. Lei, R. Teixeira, and R. Sommer, "Largescale Correlation of Accounts across Social Networks," Technical report, 2013.

[11] K. Cortis, S. Scerri, I. Rivera, and S. Handschuh, "An ontologybased technique for online profile resolution," Social Informatics, Berlin: Springer, pp. 284-298, 2013.

[12] F. Abel, E. Herder, G.J. Houben, N. Henze, and D. Krause, "Cross-system user modeling and personalization on the social web," User Modeling and User-Adapted Interaction, vol. 23, pp. 169-209, 2013.

[13] O. De Vel, A. Anderson, M. Corney, and G. Mohay, "Mining email content for author identification forensics," ACM Sigmod Record, vol. 30, no. 4, pp. 55-64, 2001.

[14] E. Raad, R. Chbeir, and A. Dipanda, "User profile matching in social networks," Proc. Of the 13th International Conference on Network-Based Information Systems (NBiS'10), pp.297-304, 2010.

[15] J. Vosecky, D. Hong, and V.Y. Shen, "User identification across multiple social networks," Proc. Of the 1st International Conference on Networked Digital Technologies, pp.360-365, 2009.

[16] P. Jain, P. Kumaraguru, and A. Joshi, "@ i seek 'fb. me': identifying users across multiple online social networks," Proc. of the 22nd International Conference on World Wide Web Companion, pp. 1259-1268, 2013

[17] P. Jain and P. Kumaraguru, "Finding Nemo: searching and resolving identities of users across online social networks," arXiv preprint arXiv:1212.6147, 2012.

[18] R. Zheng, J. Li, H. Chen, and Z. Huang, "A framework for authorship identification of online messages: writing-style fea- tures and classification techniques," J. of the American Society for Information Science and Technology, vol. 57, no. 3, pp. 378-393, 2006.

[19] M. Almishari and G. Tsudik, "Exploring linkability of user reviews," Computer Security–ESORICS 2012 (ESORICS'12), pp. 307- 324, 2012.
[20] X. Kong, J. Zhang, and P.S. Yu, "inferring anchor links across multiple heterogeneous social networks," Proc. of the 22nd ACM International Conf. on Information and Knowledge Management (CIKM'13), pp. 179-188, 2013.
[21] O. Goga, H. Lei, S.H.K. Parthasarathi, G. Friedland, R. Sommer, and R. Teixeira, "Exploiting innocuous activity for correlating users across sites," Proc. 22nd international conference on World Wide Web (WWW'13),pp. 447-458, 2013.
[22] A. Narayanan and V. Shmatikov, "De-anonymizing social networks," Proc. Of the 30th IEEE Symposium on Security and Privacy (SSP'09), pp. 173-187, 2009.
[23] S. Bartunov, A. Korshunov, S. Park, W. Ryu, and H. Lee, "Joint link-attribute user identity resolution in online social networks," The 6th SNA-KDD Workshop '12, 2012.